係数成分制約を用いた回帰分析法の提案と北海道の稲作生産性に対する気温の影響分析

姜　興　起

1．はじめに

ある確率変数 y の変動に対して，それを説明していると推測される m 個の変数 $x_1, x_2, \ldots, x_m$ があるとし，y の変動がこれらの変数の線形結合で表現される統計モデル

$$y = \beta_0 + \sum_{j=1}^{m} x_j \beta_j + \varepsilon \tag{1}$$

を線形回帰モデルという．(1) 式において，y は統計解析の対象そのものであり，目的変数という．$x_1, x_2, \ldots, x_m$ は y の変動を表現するものであり，説明変数という．これら説明変数の係数 $\beta_1, \beta_2, \ldots, \beta_m$ は回帰係数と呼ばれる．β_0 は各説明変数の変動に反応しない一定値を取る未知パラメータであり，定数項と呼ばれる．これに対して，ε は各説明変数と関係しない確率変数であり，誤差項という．明らかに，誤差項 ε が確率変数であるため，各説明変数が一定であっても，目的変数 y は確率変数である．誤差項の期待値がゼロと仮定するとき，所与の説明変数値に対して y の期待値 $\mathrm{E}\{y\} = \mu$ が

$$\mu = \beta_0 + \sum_{j=1}^{m} x_j \beta_j \tag{2}$$

で与えられる．(2) 式は回帰式と呼ばれる．

回帰分析は，予測と制御をはじめ構造解析や影響要因分析などに有効な方法として広く利用されている．その基本的な作業は次の通りである．(1) 準備作業：分析の目的に応じて目的変数と使用可能な説明変数を決定し，関連データの収集を行うこと．(2) 誤差分布の想定：誤差項の確率分布を想定すること．(3) モデルの設定：回帰式を設定すること．(4) モデルの推定：標本データを用いて設定されたモデルにおけるパラメータ値を決定すること．(5) モデルの選択：考え得るモデルが二つ以上ある場合，定められた規準に基づいてその中から一つを「最良のモデル」として選出すること．(6) モデルの応用：得られた回帰モデルを実際問題の分析に適用すること．本稿では，モデルの設定法に重点を置くことにする．本来，誤差分布の想定もモデル設定の範疇に入るべきであるが，ここでは特にそれを問題とせず，便宜上それをモデルの設定から分離した．

従来，モデル設定の問題は一般的に説明変数の選択，つまり，利用可能な説明変数のそれぞれについて各変数をモデルに導入すべきか否かを決定する問題として取り扱われている．これは，利用可能な説明変数の全部を取り込んだモデルについて一部分の回帰係数値を強制的にゼロと設定するか否かというパラメータ操作の問題としても取り扱えるから，本稿ではこれを係数制約法

(constrained coefficient approach) と呼ぶことにする．回帰分析では，モデルの推定とモデルの選択も重要な作業であるが，ここで，モデルの推定は広く利用されている最尤法を用いるので，特に問題にならない．モデルの推定と比較して，モデルの選択はよりモデルの設定に密接に関連している．従来は，自由度調整済み決定係数によるモデルチェックの方法や F 検定などの統計的検定によるモデル選択の手法が使われていた．また近年では，AIC と BIC や MDL などの情報量規準によるモデル選択の方法が主流になりつつある（坂元・石黒・北川 (1983)，小西・北川 (2004) と松嶋 (1996)）．これは，情報量規準の使用によってモデルの選択がより正確かつ容易にできるためである．しかし，係数制約法でモデルの設定を行うとき，候補となるモデルの数は説明変数の増加にしたがって急速に増大する．そのため利用可能な説明変数が多い場合，モデルの選択はモデルの設定とともに回帰分析の難問である．そのため，後退消去法や前進選定法及び段階的回帰などのモデル選択の手順が便宜的に利用されている (Draper and Smith(1981), Sen and Srivastava(1990))．また，従来のモデル設定法では，モデルから除外しようとする説明変数に含まれる目的変数の情報を完全に割愛し，情報損失の最小化に対する創意工夫が見られない．それによって，モデルの選択は膨大な候補モデル群について考えざるを得ないにも関わらず，他にもよりよいモデルが存在する可能性について判然としないままである．

本稿では，従来の係数制約法によるモデルの設定に代わって，係数成分制約法 (constrained coefficient-component approach) というモデル設定の方法を提案する．第 2 節で提案する方法についての考え方と研究の動機を述べる．第 3 節では，具体的に提案法の基礎となるモデルの独立成分化と提案法の手続きを提示し，そのパフォーマンス評価に関する理論的な考察を行う．そして第 4 節では，提案法の応用例として北海道における稲作生産性と気温との関係を分析する．

2．考え方と動機

上述のように，回帰モデルは説明変数の使い方によって決まる．回帰分析において，標本のサイズとモデルに取り込んでいる説明変数（定数項を含む）の数との差をモデルの自由度という．標本のサイズが一定のとき，モデルに取り込んだ説明変数の数が大きくなればモデルの自由度は小さくなり，モデルの標本データに対する説明力が高くなる一方，回帰係数推定値の安定性が悪くなる．従来の係数制約法では，それぞれの説明変数についてそれをモデルに取り込むか否かという二者択一の取り扱い方が採られている．そこで，モデルに取り込むべき説明変数を取り込まない場合には，モデルの自由度過剰のため，目的変数の変動が十分に説明されず，推定値と予測値に偏りが生じる．モデルの説明力の観点から言えば，多くの説明変数をモデルに取り込む方が望ましいであろう．しかし，説明変数を余分に取り込んだ場合には，モデルの自由度不足により，回帰係数推定値の安定性が悪くなるため，却って予測値のばらつきが大きくなる．推定値の安定性

の観点から，できるだけモデルに取り込む説明変数の数を制限すべきと考えられる．

このように，説明変数の利用においてモデルの説明力と推定値の安定性との間にトレードオフの関係が存在する．これを調整するために，情報量規準 AIC などが適用される（坂元・石黒・北川 (1983)）．係数制約法の考え方によると，モデルにおける説明変数の増減により，モデルの自由度を調節でき，モデルの説明力と推定値の安定性とのバランスが取れる．つまり，同様な説明力を持つモデルについて，自由度のより高いものが選択される．ところが，特定の説明変数が目的変数とある程度の相関を持つ以上，それをモデルから除外するときに必ず情報の損失が生じる．係数制約法では，このような情報損失のコントロールに対する創意工夫がまだ不十分と考えられる．また，構造解析や影響要因分析の立場から考えても，各説明変数は目的変数と関係するか否かの判断より，むしろ各説明変数の目的変数に対する影響の方向と強弱の分析がより重要である．したがって，利用可能な説明変数を簡単にモデルから除去することは必ずしも最適な方策とは言えないであろう．

本稿では，(1) 式に示すように目的変数と関係し応用上にも意味があると推測される説明変数が全部で m 個あると仮定し，これら説明変数の全部を取り込んだモデルを基本モデル（basic model）と呼ぶ．従来のモデル設定法では，ある説明変数 x_i を候補モデルに取り込もうとするとき，それに対応する回帰係数 β_i をパラメータとして取り扱い，当該説明変数を候補モデルから除外しようとするとき，$\beta_i = 0$ という制約を課す．ここで，β_i についての係数制約を

$$\mathrm{Re}(\beta_i) : \beta_i = 0 \quad (i = 1, 2, \ldots, m) \tag{3}$$

で表し，基本モデルに対してあり得る係数制約の全体集合を

$$\boldsymbol{\mathcal{B}} \equiv \{\mathrm{Re}(\beta_1), \mathrm{Re}(\beta_2), \ldots, \mathrm{Re}(\beta_m)\} \tag{4}$$

で定義する．係数制約集合 $\boldsymbol{\mathcal{B}}$ の各部分集合はそれぞれ一つの候補モデルに対応するから，候補モデルの数は，係数制約集合 $\boldsymbol{\mathcal{B}}$ の異なる部分集合の数に等しく，2^m である．

従来の係数制約法によるモデル設定のもう一つの問題は，利用可能な説明変数が多いとき，正確なモデル選択が困難とのことである．もし AIC などの基準によって係数制約集合 $\boldsymbol{\mathcal{B}}$ の係数制約に関する順序関係を定義できれば，考え得る候補モデルが $m+1$ 個だけあり，モデル選択の作業がより単純になる．しかし，一般的に各説明変数は互いに影響し合うので，ある回帰係数に関する制約の効果は他の回帰係数に関する制約の使い方によって異なる．そのため，係数制約集合 $\boldsymbol{\mathcal{B}}$ では係数制約の順序関係を定義することが不可能となり，$\boldsymbol{\mathcal{B}}$ の全ての部分集合についてモデルの選択を考えざるを得ない．このように，基本モデルにおける説明変数の数 m の増大につれて，候補モデルの数は指数関数 2^m で増大する．例えば，$m = 10$ のとき，候補モデルの数は $2^{10} = 1024$ となってしまう．これほど膨大な候補モデル群においてモデルの選択を行うことは必ずしも容易

なことではない．この問題に対処するために，鈴木・後藤・俵 (2000) は，ベイズ法を用いた回帰モデルによる予測の方法を提案しているが，簡易性と汎用性を持つアルゴリズムの開発は課題として残っている．また，その方法では予測信頼区間の構成が困難であり，構造解析へ適用し難いなどの問題もある．

本稿では，従来の係数制約法に代わって係数成分制約法を提案する．提案法では，直交変換によって基本モデルを幾つかの独立な部分モデルに変換し，変換後のモデルは回帰係数に代わってその線形結合である係数成分をパラメータとして持つ．各係数成分がそれぞれ独立な部分モデルにあるから，それぞれに関する操作を独立に行うことができる．そのため，モデルの設定は各係数成分の組み合わせで考える必要がなくなる．これによって考え得る候補モデルの数が減少し，モデルの選択がより容易になるだけでなく，パラメータ推定値の改良やよりよいモデル発見の可能性が得られる．具体的に提案法は次のようなメリットを持つ．(1) パラメータ推定値の改良とよりよいモデルの発見が可能である．(2) 考え得る候補モデルの数を最小限に抑えることによって，モデルの選択をより容易に実現できる．(3) アルゴリズムの簡易性と汎用性を追求できる．

3．係数成分制約法の提案

3.1 標本モデル

前述のように本稿では，(1) 式のモデルを基本モデルとして考える．ここで，$n+1$ 組の標本データ $y_0, y_1, \ldots, y_n$ についての基本モデルを

$$y_i = \beta_0 + \sum_{j=1}^{m} x_{ij}\beta_j + \varepsilon_i \quad (i = 0, 1, \ldots, n) \tag{5}$$

で表現する．ただし，$x_{i1}, x_{i2}, \ldots, x_{im}$ と ε_i はそれぞれ i 番目の標本値 y_i に対応する説明変数と誤差項である．一般に，(5) 式のモデルについて次のような仮定を置く．(1) 各説明変数値は所与の定数である．(2) 各誤差項は期待値 0，共通の分散 σ^2（未知）をもつ確率変数である．(3) 各誤差項は互いに独立である．(4) 全ての誤差項は正規分布に従う．

仮定の (1) は問題の単純化のために一般に用いられている．仮定の (2) と (3) は，回帰係数の最小二乗推定量（least squares estimator）が最良線形不偏推定量（BLUE: best linear unbiased estimator）になることを保証し，Gauss-Markov の条件と呼ばれる (Sen and Srivastava(1990))．また，仮定の (4) の下で，パラメータの最小二乗推定量が最尤推定量（maximum likelihood estimator）に一致することが広く知られている．因みに，(5) 式の基本モデルにおいて，定数項 β_0 と回帰係数 $\beta_1, \beta_2, \ldots, \beta_m$ 及び誤差分散 σ^2 が未知パラメータとなり，モデルの自由度が $n-m$ である．一般的に意味のあるパラメータの推定を得るには，自由度 $n-m$ が 1 を超えること，つまり，$n > m+1$ になることが必要である．

3.2 モデルの中心化と行列・ベクトル表現

モデルの中心化とは各説明変数の平均がゼロになるように

$$\widetilde{x}_{ij} = x_{ij} - \bar{x}_j \quad (j = 1, 2, \ldots, m) \tag{6}$$

で変数変換を行い，変換後の $\widetilde{x}_{ij}$ を新しい説明変数として用いることである．ただし，$\bar{x}_j$ は j 番目の説明変数の平均を表し，次式で与えられる．

$$\bar{x}_j = \frac{1}{n+1}\sum_{i=0}^{n} x_{ij}$$

また，(2) 式により，y_i に対応する回帰式は

$$\mu_i = \beta_0 + \sum_{j=1}^{m} x_{ij}\beta_j$$

で与えられる．さらに，$\mu_0, \mu_1, \ldots, \mu_n$ の平均を α_0 で表せば，

$$\alpha_0 = \frac{1}{n+1}\sum_{i=0}^{n} \mu_i = \beta_0 + \sum_{j=1}^{m} \bar{x}_j\beta_j \tag{7}$$

となる．ここで，α_0 を（目的変数の）平均期待という．(6) 式と (7) 式により，(5) 式の基本モデルは次のようになる．

$$y_i = \alpha_0 + \sum_{j=1}^{m} \widetilde{x}_{ij}\beta_j + \varepsilon_i \quad (i = 0, 1, \ldots, n) \tag{8}$$

また，最尤法あるいは最小二乗法でパラメータを推定する際に，定数項の推定に先行して回帰係数の推定を行うことができる．それは各回帰係数の推定値が目的変数と各説明変数との共分散で先に決まり，定数項の推定値が回帰係数の推定値と目的変数の平均によって定義されるからである（Sen and Srivastava(1990)）．(7) 式から解るように，各回帰係数が推定されたとき，平均期待 α_0 と定数項 β_0 との推定値の間に一対一の対応関係がある．したがって，(8) 式の中心化されたモデルは (5) 式の基本モデルと同値である．明らかに，平均期待 α_0 は中心化されたモデルにおける定数項と見なすことができる．

行列・ベクトルの表現により，(8) 式のモデルは次のように表される．

$$\boldsymbol{y} = \boldsymbol{X}\begin{bmatrix} \alpha_0 \\ \boldsymbol{\beta} \end{bmatrix} + \boldsymbol{\varepsilon} \tag{9}$$

ここで，$\boldsymbol{y} = (y_0, y_1, \ldots, y_n)^{\mathrm{t}}$ は $n+1$ 次の目的変数ベクトルで，$\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_m)^{\mathrm{t}}$ は m 次の回帰係数ベクトルであり，$\boldsymbol{\varepsilon} = (\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n)^{\mathrm{t}}$ は誤差ベクトルである．$\boldsymbol{X}$ は $(n+1) \times (m+1)$ のデザイン行列であり，次のように定義される．

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{1}_{n+1} & \widetilde{\boldsymbol{X}} \end{bmatrix}$$

ただし，$\widetilde{\boldsymbol{X}} = (\widetilde{x}_{ij})$ は各中心化された説明変数値で定義される $(n+1) \times m$ の説明変数行列を表し，$\mathbf{1}_{n+1}$ は全ての要素が 1 の $n+1$ 次のベクトルを表す．説明変数行列 $\widetilde{\boldsymbol{X}}$ の各列における要素の平均値が常に 0 であるため，

$$\widetilde{\boldsymbol{X}}^{\mathrm{t}} \mathbf{1}_{n+1} = \mathbf{0}_m \tag{10}$$

という関係が得られる．ただし，$\mathbf{0}_m$ は全ての要素が 0 の m 次のベクトルを表す．

(5) 式のモデルに関する仮定の (2)，(3) と (4) から解るように，誤差ベクトル $\boldsymbol{\varepsilon}$ は期待値が $\mathbf{0}_{n+1}$，共分散行列が $\sigma^2 \boldsymbol{I}_{n+1}$ の $n+1$ 変量正規分布に従う．統計学の表現を用いると，

$$\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}_{n+1}, \sigma^2 \boldsymbol{I}_{n+1}) \tag{11}$$

となる．ただし，$\boldsymbol{I}_{n+1}$ は $n+1$ 次の単位行列を表している．また，本稿では多重共線性のことを問題にしないため，デザイン行列 $\boldsymbol{X}$ はランクが $m+1$ のフルランク行列であると仮定する．

3.3 モデルの独立成分化

(9) 式における基本モデルの構造を保持しながら回帰係数などのパラメータに関わる部分モデルの次元を縮約するために，次のような Householder 変換を行う．

$$\boldsymbol{H}\boldsymbol{y} = \boldsymbol{H}\boldsymbol{X} \begin{bmatrix} \alpha_0 \\ \boldsymbol{\beta} \end{bmatrix} + \boldsymbol{H}\boldsymbol{\varepsilon} \tag{12}$$

ただし，$\boldsymbol{H}$ は $n+1$ 次の直交行列である．

鏡映変換を用いた Householder 変換の詳細なアルゴリズムは坂元・石黒・北川 (1983) と北川 (1993, 2005) にあるが，本稿では次のようにそれを適用する．まず，適当に設定した直交行列 $\boldsymbol{H}$ を用いて，デザイン行列 $\boldsymbol{X}$ を

$$\boldsymbol{H}\boldsymbol{X} = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{O} \end{bmatrix} \tag{13}$$

のように変換し，$\boldsymbol{H}\boldsymbol{y}$ の計算を行う．(13) 式において，$\boldsymbol{A}$ は $m+1$ 次の上三角行列であり，$\boldsymbol{O}$ は全ての要素が 0 の零行列である．行列 $\boldsymbol{A}$ の詳細な構造は次の通りである．

$$\boldsymbol{A} = \begin{bmatrix} d_0 & \boldsymbol{g}^{\mathrm{t}} \\ \mathbf{0}_m & \boldsymbol{T} \end{bmatrix} \tag{14}$$

ここで，d_0 と $\boldsymbol{g}$ はそれぞれスカラーと m 次のベクトルであり，$\boldsymbol{T}$ は m 次の上三角行列である．直交行列の性質により，(13) 式から

$$\boldsymbol{X}^{\mathrm{t}}\boldsymbol{X} = \boldsymbol{X}^{\mathrm{t}}\boldsymbol{H}^{\mathrm{t}}\boldsymbol{H}\boldsymbol{X} = \boldsymbol{A}^{\mathrm{t}}\boldsymbol{A}$$

という関係を得る．よって，$d_0^2 = \mathbf{1}_{n+1}^{\mathrm{t}} \mathbf{1}_{n+1} = n+1$ と $d_0 \boldsymbol{g} = \widetilde{\boldsymbol{X}}^{\mathrm{t}} \mathbf{1}_{n+1}$ とのことを容易に確認できる．さらに，(10) 式の関係を考慮すると，$d_0 \neq 0$ と $\boldsymbol{g} = \mathbf{0}_m$ が必然な結果として得られる．また，デザイン行列 $\boldsymbol{X}$ のランクが $m+1$ であるため，行列 $\boldsymbol{T}$ は非特異な行列となる．

次に，前述の Householder 変換に用いられる直交行列 $\boldsymbol{H}$ を $\boldsymbol{H}^{\mathrm{t}} = \left[\begin{array}{ccc} \boldsymbol{H}_0^{\mathrm{t}} & \boldsymbol{H}_1^{\mathrm{t}} & \boldsymbol{H}_2^{\mathrm{t}} \end{array}\right]$ のように分割する．ここで，$\boldsymbol{H}_0$，$\boldsymbol{H}_1$ と $\boldsymbol{H}_2$ はそれぞれ $1 \times (n+1)$，$m \times (n+1)$ と $(n-m) \times (n+1)$ 次の行列である．このように，Householder 変換の結果を用いると，(9) 式の基本モデルは次のようになる．

$$
\begin{aligned}
z_0 &= d_0\alpha_0 + \boldsymbol{H}_0\boldsymbol{\varepsilon} && (15)\\
\boldsymbol{w} &= \boldsymbol{T}\boldsymbol{\beta} + \boldsymbol{\eta} && (16)\\
\boldsymbol{e} &= \boldsymbol{H}_2\boldsymbol{\varepsilon} && (17)
\end{aligned}
$$

ここで，

$$
\begin{aligned}
z_0 &= \boldsymbol{H}_0\boldsymbol{y} &&\\
\boldsymbol{w} &= (w_1, w_2, \ldots, w_m)^{\mathrm{t}} = \boldsymbol{H}_1\boldsymbol{y} && (18)\\
\boldsymbol{\eta} &= (\eta_1, \eta_2, \ldots, \eta_m)^{\mathrm{t}} = \boldsymbol{H}_1\boldsymbol{\varepsilon} &&\\
\boldsymbol{e} &= (e_1, e_2, \ldots, e_{n-m})^{\mathrm{t}} = \boldsymbol{H}_2\boldsymbol{y} && (19)
\end{aligned}
$$

と定義される．(11) 式と直交行列の性質から解るように，$\boldsymbol{H}_0\boldsymbol{\varepsilon}$，$\boldsymbol{\eta}$ と $\boldsymbol{H}_2\boldsymbol{\varepsilon}$ は互いに独立であり，それぞれの分布は次のようになる．

$$
\begin{aligned}
\boldsymbol{H}_0\boldsymbol{\varepsilon} &\sim \mathrm{N}(0, \sigma^2) && (20)\\
\boldsymbol{\eta} &\sim \mathrm{N}(\boldsymbol{0}_m, \sigma^2\boldsymbol{I}_m) && (21)\\
\boldsymbol{H}_2\boldsymbol{\varepsilon} &\sim \mathrm{N}(\boldsymbol{0}_{n-m}, \sigma^2\boldsymbol{I}_{n-m}) && (22)
\end{aligned}
$$

このようにして，平均期待 α_0 と回帰係数ベクトル $\boldsymbol{\beta}$ に関わる部分モデルの次元が最小化され，それぞれに関わる部分モデルが互いに独立となる．

ところが，一般的に (16) 式のモデルにおける行列 $\boldsymbol{T}$ が対角行列でないため，各回帰係数が互いに関連して，それぞれに関する操作を独立に行うことはできない．そのため，次のように行列 $\boldsymbol{T}$ の特異値分解を行う．

$$
\boldsymbol{T} = \boldsymbol{U}^{\mathrm{t}}\boldsymbol{D}\boldsymbol{V} \qquad (23)
$$

(23) 式において，$\boldsymbol{U}$ と $\boldsymbol{V}$ はいずれも m 次の直交行列である．また，$\boldsymbol{D} = \mathrm{diag}(d_1, d_2, \ldots, d_m)$ は m 次の対角行列であり，その対角要素 $d_1, d_2, \ldots, d_m$ が行列 $\boldsymbol{T}$ の特異値となる．前述のことから解るように，行列 $\boldsymbol{T}$ が非特異であるため，行列 $\boldsymbol{D}$ も非特異である．したがって，$d_i \neq 0\ (i = 1, 2, \ldots, m)$ が保証される．因みに，一般的なアルゴリズムでは，各特異値について $d_i > 0$ という条件が課される．行列 $\boldsymbol{U}$ と $\boldsymbol{V}$ の設定を工夫すればこれを保証できるが，本研究においてはこれに固執しな

い．一般的な特異値分解のアルゴリズムは Golub and Loan(1989) にあるが，本稿では，行列 $\boldsymbol{T}$ が非特異な正方行列であることを利用して，その特異値分解を次のような固有値問題にして解くことにする．

さて，(23) 式により，

$$\boldsymbol{T}^{\mathrm{t}}\boldsymbol{T} \;=\; \boldsymbol{V}^{\mathrm{t}}\boldsymbol{D}^2\boldsymbol{V} \tag{24}$$

$$\boldsymbol{T}\boldsymbol{T}^{\mathrm{t}} \;=\; \boldsymbol{U}^{\mathrm{t}}\boldsymbol{D}^2\boldsymbol{U} \tag{25}$$

という関係が得られる．(24) 式と (25) 式において，行列 $\boldsymbol{T}^{\mathrm{t}}\boldsymbol{T}$ と $\boldsymbol{T}\boldsymbol{T}^{\mathrm{t}}$ はいずれも m 次の正値対称行列であるから，それぞれ m 個の正の実固有値を持つ．また，(24) 式と (25) 式から $\boldsymbol{T}^{\mathrm{t}}\boldsymbol{T}$ と $\boldsymbol{T}\boldsymbol{T}^{\mathrm{t}}$ は同様に $d_1^2, d_2^2, \ldots, d_m^2$ を固有値として持つことが解る．これらの固有値が互いに異なる場合，直交行列 $\boldsymbol{V}$ と $\boldsymbol{U}$ はそれぞれ (24) 式と (25) 式で定義される固有値問題によって一意に求まる．(23) 式を書き換えると，$\boldsymbol{D} = \boldsymbol{U}\boldsymbol{T}\boldsymbol{V}^{\mathrm{t}}$ となり，つまり，各特異値は次のように求められる．

$$d_i = \boldsymbol{u}_i^{\mathrm{t}}\boldsymbol{T}\boldsymbol{v}_i \quad (i = 1, 2, \ldots, m) \tag{26}$$

ただし，$\boldsymbol{u}_i$ と $\boldsymbol{v}_i$ はそれぞれ行列 $\boldsymbol{U}^{\mathrm{t}}$ と $\boldsymbol{V}^{\mathrm{t}}$ の第 i 列のベクトルを表す．

ここで，回帰係数ベクトル $\boldsymbol{\beta}$ の直交変換

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m)^{\mathrm{t}} = \boldsymbol{V}\boldsymbol{\beta} \tag{27}$$

で m 次のベクトル $\boldsymbol{\alpha}$ を定義し，さらに，

$$\boldsymbol{z} \;=\; (z_1, z_2, \ldots, z_m)^{\mathrm{t}} = \boldsymbol{U}\boldsymbol{w} \tag{28}$$

$$\boldsymbol{\psi} \;=\; (\psi_1, \psi_2, \ldots, \psi_m)^{\mathrm{t}} = \boldsymbol{U}\boldsymbol{\eta} \tag{29}$$

のような直交変換を用いると，(16) 式のモデルは

$$\boldsymbol{z} = \boldsymbol{D}\boldsymbol{\alpha} + \boldsymbol{\psi} \tag{30}$$

と表される．このとき，$\boldsymbol{\alpha}$ の各要素は基本モデルにおける回帰係数の線形結合で定義されるものであり，それぞれを係数成分と呼び，$\boldsymbol{\alpha}$ を係数成分ベクトルという．

以上みてきたように，(9) 式の基本モデルは (15) 式，(30) 式と (17) 式で表され，その中で特に (30) 式のモデルは重要な役割を果たすものである．(30) 式のモデルにおけるベクトル $\boldsymbol{z}$ がモデルを決める要であり，それを枢軸ベクトルといい，その各要素は枢軸成分と呼ばれる．明らかに，(29) 式で定義されるベクトル $\boldsymbol{\psi}$ は $\boldsymbol{H}_0\boldsymbol{\varepsilon}$ 及び $\boldsymbol{H}_0\boldsymbol{\varepsilon}$ と独立である．また，(21) 式と (29) 式により，$\boldsymbol{\psi}$ の分布は次のように表される．

$$\boldsymbol{\psi} \sim \mathrm{N}(\boldsymbol{0}_m, \sigma^2\boldsymbol{I}_m) \tag{31}$$

ここで，(30) 式のモデルを成分ごとに表現すると，

$$z_i = d_i\alpha_i + \psi_i \quad (i = 1, 2, \ldots, m) \tag{32}$$

となる．(32) 式のモデルから，次のことが解る．$i = 1, 2, \ldots, m$ について，それぞれの枢軸成分に関するモデルに一つの係数成分があり，また，各枢軸成分が互いに独立である．そのため，それぞれの係数成分に関する操作を独立に行うことができる．

3.4 係数成分制約法の手続き

ここで，モデル設定の新しい方法として係数成分制約法を提案する．この方法では，回帰係数に代わって (30) 式または (32) 式のモデルにおける各係数成分に関する操作によってモデルの設定を行う．

本稿では，簡単のため平均期待 α_0 を未知パラメータとして取り扱い，その最尤推定値を次式によって求める．

$$\widehat{\alpha}_0 = \frac{z_0}{d_0} \tag{33}$$

係数制約法と同様に，係数成分についての操作は基本的に 2 通りある．一つは係数成分にゼロ制約を付加することであり，もう一つはそれを未知パラメータとして取り扱うことである．具体的に，$i = 1, 2, \ldots, m$ について，前者は係数成分 α_i について

$$\alpha_i = 0 \tag{34}$$

でそれを制約することであり，後者はそれを未知パラメータとして最尤推定によって

$$\widehat{\alpha}_i = \frac{z_i}{d_i} \tag{35}$$

で推定することである．従来の係数制約法では，特別な回帰係数をゼロにすることによってモデルの自由度を調節するが，これは説明変数選択の観点から解釈すれば極めて自然なことである．しかし，係数成分について自由度調節の観点から見れば，ゼロ以外の定数で制約を施すことも考えられるから，(34) 式のようにゼロで制約することの意味は必ずしも自明ではない．これは予測値の安定性の観点から解釈できる．つまり，他の定数で関係する係数成分を制約するより，ゼロでそれを制約することは目的変数予測値の分散を最小にすることができる．本稿では，このようなモデル設定のルールを最小分散予測の原則と呼ぶことにする（付録 A を参照）．

いま，(3) 式の表現を利用すれば，(34) 式の係数成分制約は

$$\mathrm{Re}(\alpha_i) : \alpha_i = 0 \quad (i = 1, 2, \ldots, m) \tag{36}$$

で表現される．さらに，(30) 式のモデルに適用し得る係数成分制約の全体集合は

$$\boldsymbol{\mathcal{A}} \equiv \{\mathrm{Re}(\alpha_1), \mathrm{Re}(\alpha_2), \dots, \mathrm{Re}(\alpha_m)\} \tag{37}$$

によって定義され，(30) 式のモデルに係数成分制約集合 $\boldsymbol{\mathcal{A}}$ の特定な部分集合における係数成分制約を適用すれば，相応の自由度を持つ候補モデルが設定される．前述のように設定できるモデルは全部で 2^m 個もあるが，もし AIC などの規準で係数成分制約集合 $\boldsymbol{\mathcal{A}}$ における各係数成分制約についての優先順位を定義できれば，競争に勝つことのできないものは事前に候補モデルから排除される．これによって，考え得る候補モデルの数が減少し，モデル選択作業の単純化することができる．そこで，係数成分制約集合 $\boldsymbol{\mathcal{A}}$ における各係数成分制約の優先順位の定義と方法が問題となる．これについて，次の定理が導かれる（証明は付録 B を参照）．

定理 1 (32) 式のモデルにおける各枢軸成分絶対値の全体集合 $\{|z_1|, |z_2|, \dots, |z_m|\}$ において，互いに等しい元が存在しないと仮定する．$|z_1|, |z_2|, \dots, |z_m|$ の大きい順で $z_1, z_2, \dots, z_m$ の系列を並べ替えたものを

$$z_1^*, z_2^*, \dots, z_m^* \tag{38}$$

で表し，これに対応して係数成分 $\alpha_1, \alpha_2, \dots, \alpha_m$ の系列を並べ替えたものを

$$\alpha_1^*, \alpha_2^*, \dots, \alpha_m^* \tag{39}$$

とする．また，$0 \le k \le m$ を満たす整数 k について $\boldsymbol{\mathcal{A}}$ の部分集合 $\boldsymbol{\mathcal{A}}_k$ を，$k > 0$ ときに

$$\boldsymbol{\mathcal{A}}_k \equiv \{Re(\alpha_{m-k+1}^*), Re(\alpha_{m-1}^*), \dots, Re(\alpha_m^*)\} \tag{40}$$

で，$k = 0$ のときに $\boldsymbol{\mathcal{A}}_0 = \Phi$ で定義する．ただし，Φ は空集合を表す．平均期待 α_0 を未知パラメータとしたとき，(15) 式，(30) 式と (17) 式で定義される基本モデルについて，次の命題が真である．係数成分制約集合 $\boldsymbol{\mathcal{A}}$ における任意の k 個の係数成分制約を適用したモデルの集合において，$\boldsymbol{\mathcal{A}}_k$ における係数成分制約が適用されたモデルは最小 AIC の値を与える．対応する AIC の値は

$$AIC(\boldsymbol{\mathcal{A}}_k) = (n+1)\ln\{\hat{\sigma}^2(\boldsymbol{\mathcal{A}}_k)\} + \lambda \tag{41}$$

で与えられる．ここで，λ は

$$\lambda = (n+1)(\ln\{2\pi\} + 1) + 2(m - k + 2) \tag{42}$$

で定義される定数であり，$\hat{\sigma}^2(\boldsymbol{\mathcal{A}}_k)$ は誤差分散の最尤推定を表し，次のように与えられる．

$$\begin{aligned} \hat{\sigma}^2(\boldsymbol{\mathcal{A}}_k) &= \frac{1}{n+1}\Big[\sum_{i=m-k+1}^{m} (z_i^*)^2 + \boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}\Big] \\ &= \frac{1}{n+1}\Big[\sum_{i=m-k+1}^{m} (z_i^*)^2 + \sum_{i=1}^{n-m} e_i^2\Big] \end{aligned} \tag{43}$$

ただし，$\boldsymbol{e} = (e_1, e_2, \ldots, e_{n-m})^{\mathrm{t}}$ は (19) 式で定義されている．

松嶋 (1996) と小西・北川 (2004) などが示すように，線形回帰モデルに対して，AIC のほかに BIC など多くの情報量規準も対数変換した誤差分散最尤推定の線形関数で表現される．定理 1 は，このように定義される全ての情報量規準に適用できる極めて一般性を持つ結果である．

定理 1 は，適用される係数成分制約数 k が一定のとき考え得る候補モデルが一意に決まることを示唆している．つまり，基本モデルに集合 $\boldsymbol{\mathcal{A}}_k$ における係数成分制約を適用したモデルを $\mathrm{M}(\boldsymbol{\mathcal{A}}_k)$ とすると，提案の係数成分制約法によるモデル設定の場合，モデルの選択はモデル集合 $\{\mathrm{M}(\boldsymbol{\mathcal{A}}_0), \mathrm{M}(\boldsymbol{\mathcal{A}}_1), \ldots, \mathrm{M}(\boldsymbol{\mathcal{A}}_m)\}$ について行えばよい．ここで，$\mathrm{M}(\boldsymbol{\mathcal{A}}_0)$ と $\mathrm{M}(\boldsymbol{\mathcal{A}}_m)$ は二つの極端なケースに対応していることに注意しよう．前者は基本モデルそのものである．これに対して，後者は係数成分制約集合 $\boldsymbol{\mathcal{A}}$ における全ての係数成分制約を適用したモデルである．結局これは，係数制約集合 $\boldsymbol{\mathcal{B}}$ における全ての係数制約を適用するモデル，つまり，説明変数を一つも取り込まないモデルそのものに他ならない．

上述のように最小 AIC 法を適用すれば，モデル集合 $\{\mathrm{M}(\boldsymbol{\mathcal{A}}_0), \mathrm{M}(\boldsymbol{\mathcal{A}}_1), \ldots, \mathrm{M}(\boldsymbol{\mathcal{A}}_m)\}$ の中から「最良のモデル」が簡単に選出される．また，最良のモデルについて，定理 1 を適用すると適用される係数成分制約が決まるので，これら係数成分の値をゼロと設定し，他の係数成分は最尤法によって (35) 式で推定される．つまり，最小 AIC 法と最尤法を同時に利用することによって各係数成分の推定値を決めることができる．ここで，前述の二つの極端なケースを除いて，$0 < k < m$ という範囲における一定の k の値について考える．もし，考え得るモデルの中で $\mathrm{M}(\boldsymbol{\mathcal{A}}_k)$ が最良のモデルとならば，定理 1 に用いられる表現により最尤法で推定すべき係数成分の集合は $\{\alpha_1^*, \alpha_2^*, \ldots, \alpha_{m-k}^*\}$ となり，制約を適用すべき係数成分の集合は $\{\alpha_{m-k+1}^*, \alpha_{m-k+2}^*, \ldots, \alpha_m^*\}$ となる．これに対応する係数成分ベクトルを $\boldsymbol{\alpha}^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_m^*)^{\mathrm{t}}$ で表すと，係数成分制約法により $\boldsymbol{\alpha}^*$ の推定値 $\widehat{\boldsymbol{\alpha}}^*$ は $\widehat{\boldsymbol{\alpha}}^* = (\widehat{\alpha}_1^*, \widehat{\alpha}_2^*, \ldots, \widehat{\alpha}_{m-k}^*, 0, \ldots, 0)^{\mathrm{t}}$ となる．さらに，対応する回帰係数ベクトル $\boldsymbol{\beta}$ の推定値は (27) 式の係数成分ベクトルと回帰係数ベクトルとの対応関係によって求まる．また，必要に応じて (7) 式を適用すれば，定数項 β_0 の推定値は回帰係数の推定値と平均期待 α_0 の推定値から導出可能である．

ところで，基本的には平均期待 α_0 も一つの係数成分と見なすことができる．つまり，それについて特別な制約を課すか，あるいは前述のように未知パラメータとして最尤法で推定するかというモデル設定の問題が残っている．これも最小 AIC 法などを用いれば解決できる．

3.5 係数成分制約法の評価

前述で提案した係数成分制約法により，優位になれない候補モデルを事前に排除し，モデル選択作業の単純化を実現することが可能である．これは提案法の大きなメリットである．そこで次

に，提案法に対するパフォーマンス評価，つまり，従来法と比較して提案法ではよりよいモデルを発見することが可能かどうかを検討する．

この問題を考える前の予備的考察として，従来法と提案法との共通点及び相違点について整理しておこう．従来法は (15) 式，(16) 式と (17) 式の基本モデルをベースに，モデルの設定が各回帰係数に対する制約の運用によって実現される．これに対して，提案法は (15) 式，(30) 式と (17) 式の基本モデルをベースに，モデルの設定が各係数成分に対する制約を通して行われる．(30) 式のモデルと (16) 式のモデルは，直交変換によって一方からもう一方が得られるため，互いに同値である．したがって，従来法と提案法は同値な基本モデルをベースに，それぞれ回帰係数と係数成分に対する制約を通してモデルの自由度を調節し，モデルの設定を行うのである．この二つの方法における相違点は (16) 式のモデルと (30) 式のモデルに対する操作に現れる．

ここで，(16) 式と (30) 式における m 次のベクトル $\boldsymbol{w}=(\boldsymbol{w}_1^{\mathrm{t}},\boldsymbol{w}_2^{\mathrm{t}})^{\mathrm{t}}$，$\boldsymbol{\beta}=(\boldsymbol{\beta}_1^{\mathrm{t}},\boldsymbol{\beta}_2^{\mathrm{t}})^{\mathrm{t}}$，$\boldsymbol{\eta}=(\boldsymbol{\eta}_1^{\mathrm{t}},\boldsymbol{\eta}_2^{\mathrm{t}})^{\mathrm{t}}$，$\boldsymbol{z}=(\boldsymbol{z}_1^{\mathrm{t}},\boldsymbol{z}_2^{\mathrm{t}})^{\mathrm{t}}$，$\boldsymbol{\alpha}=(\boldsymbol{\alpha}_1^{\mathrm{t}},\boldsymbol{\alpha}_2^{\mathrm{t}})^{\mathrm{t}}$ と $\boldsymbol{\psi}=(\boldsymbol{\psi}_1^{\mathrm{t}},\boldsymbol{\psi}_2^{\mathrm{t}})^{\mathrm{t}}$ のように分割し，これらに対応して各モデルにおける上三角行列 $\boldsymbol{T}$，対角行列 $\boldsymbol{D}$ 及び直交行列 $\boldsymbol{V}$ を

$$\boldsymbol{T}=\begin{bmatrix}\boldsymbol{T}_{11} & \boldsymbol{T}_{12}\\ \boldsymbol{O} & \boldsymbol{T}_{22}\end{bmatrix},\quad \boldsymbol{D}=\begin{bmatrix}\boldsymbol{D}_1 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{D}_2\end{bmatrix},\quad \boldsymbol{V}=\begin{bmatrix}\boldsymbol{V}_{11} & \boldsymbol{V}_{12}\\ \boldsymbol{V}_{21} & \boldsymbol{V}_{22}\end{bmatrix}$$

のようにブロック化する．ここで，$\boldsymbol{w}_2$，$\boldsymbol{\beta}_2$，$\boldsymbol{\eta}_2$，$\boldsymbol{z}_2$，$\boldsymbol{\alpha}_2$ と $\boldsymbol{\psi}_2$ を k 次のベクトル，$\boldsymbol{T}_{22}$，$\boldsymbol{D}_2$ と $\boldsymbol{V}_{22}$ を k 次の正方行列とすると，他の部分ベクトル及び部分行列の次元が明らかになる．ただしここでは，$0<k<m$ について考える．これらの表現を用いると，(16) 式のモデルは

$$\begin{bmatrix}\boldsymbol{w}_1\\ \boldsymbol{w}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{T}_{11} & \boldsymbol{T}_{12}\\ \boldsymbol{O} & \boldsymbol{T}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{bmatrix}+\begin{bmatrix}\boldsymbol{\eta}_1\\ \boldsymbol{\eta}_2\end{bmatrix} \tag{44}$$

となり，(30) 式のモデルは

$$\begin{bmatrix}\boldsymbol{z}_1\\ \boldsymbol{z}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{D}_1 & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{D}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\end{bmatrix}+\begin{bmatrix}\boldsymbol{\psi}_1\\ \boldsymbol{\psi}_2\end{bmatrix} \tag{45}$$

となる．また，(27) 式は

$$\begin{bmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{V}_{11} & \boldsymbol{V}_{12}\\ \boldsymbol{V}_{21} & \boldsymbol{V}_{22}\end{bmatrix}\begin{bmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{bmatrix} \tag{46}$$

で表される．

明らかに，回帰係数ベクトル $\boldsymbol{\beta}$ と係数成分ベクトル $\boldsymbol{\alpha}$ の推定値はそれぞれ (44) 式の部分モデルと (45) 式のモデルによって決まる．次の補題は容易に得られるが，ここで，後続結果の基礎としてこれを提示しておく（証明は省略）．

補題 1 基本モデルの条件下では，$\boldsymbol{\alpha}_1$ と $\boldsymbol{\alpha}_2$ の最尤推定値 $\widehat{\boldsymbol{\alpha}}_1$ と $\widehat{\boldsymbol{\alpha}}_2$ はそれぞれ $\widehat{\boldsymbol{\alpha}}_1=\boldsymbol{D}_1^{-1}\boldsymbol{z}_1$ と $\widehat{\boldsymbol{\alpha}}_2=\boldsymbol{D}_2^{-1}\boldsymbol{z}_2$ で与えられる．また，誤差分散 σ^2 が与えられたとき，$\widehat{\boldsymbol{\alpha}}_1\sim N(\boldsymbol{\alpha}_1,\sigma^2\boldsymbol{D}_1^{-2})$ と $\widehat{\boldsymbol{\alpha}}_2\sim N(\boldsymbol{\alpha}_2,\sigma^2\boldsymbol{D}_2^{-2})$ は互いに独立である．

従来法によるモデル設定について一般性を失うことなく，

$$\boldsymbol{\beta}_2 = \mathbf{0}_k \tag{47}$$

という係数制約をおく．(44) 式のモデルに (47) 式の係数制約を適用すると

$$\begin{bmatrix} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{T}_{11} \\ \boldsymbol{O} \end{bmatrix} \boldsymbol{\beta}_1 + \begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \end{bmatrix} \tag{48}$$

となる．一方，これに対応して提案法によるモデル設定について

$$\boldsymbol{\alpha}_2 = \mathbf{0}_k \tag{49}$$

という係数成分制約をおく．(45) 式のモデルにこれを用いると

$$\begin{bmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{D}_1 \\ \boldsymbol{O} \end{bmatrix} \boldsymbol{\alpha}_1 + \begin{bmatrix} \boldsymbol{\psi}_1 \\ \boldsymbol{\psi}_2 \end{bmatrix} \tag{50}$$

となる．このように，係数制約法では，(15) 式，(48) 式と (17) 式で表現されるモデルが得られ，係数成分制約法では，(15) 式，(50) 式と (17) 式で表現されるモデルがある．この 2 種類のモデル設定について，次の補題が導出される（証明は付録 C を参照）．

補題 2 基本モデルと最小分散予測の原則の下では，(47) 式の係数制約が成り立つならば，(49) 式の係数成分制約は必ず成り立つ．つまり，前者は後者の十分条件である．

補題 2 を適用すれば次の定理が得られる（証明は付録 D を参照）．

定理 2 係数制約法で設定された (15) 式，(48) 式と (17) 式のモデルに関する誤差分散の最尤推定を $\hat{\sigma}_1^2(k)$ とし，係数制約法で設定された (15) 式，(50) 式と (17) 式のモデルに関する誤差分散の最尤推定を $\hat{\sigma}_2^2(k)$ とする．(47) 式の係数制約と (49) 式の係数成分制約の少なくとも一つが成り立つならば，次の不等式が成り立つ．

$$E\{\hat{\sigma}_1^2(k)\} \geq E\{\hat{\sigma}_2^2(k)\} \tag{51}$$

ただし，E は基本モデルに関する期待値を表す．

定理 2 は，前述 2 種類モデルの少なくとも一つが真なモデルに近ければ，平均的に提案した係数成分制約法では従来の係数制約法より小さい誤差分散の推定が得られることを示している．また，AIC の定義により，(51) 式が成り立てば次の関係が成立することが解る．

$$\mathrm{E}\Big[\exp\{AIC_1(k)\}\Big] \geq \mathrm{E}\Big[\exp\{AIC_2(k)\}\Big] \tag{52}$$

ここで，E は (51) 式のそれと同様な意味を持つ．また，$AIC_1(k)$ と $AIC_2(k)$ はそれぞれ $\hat{\sigma}_1^2(k)$ と $\hat{\sigma}_2^2(k)$ に対応する AIC の値を表す．実際の応用では，設定したモデルは真のモデルに近いか否かは最小 AIC 法に基づいて経験的に判別できる．特に，定理1を適用すれば真のモデルに近いモデルを最良のモデルとして特定することができる．さらに，定理2により，提案の係数成分制約法で得られた最良のモデルは従来の係数制約法で得られる「最良のモデル」よりもよいパフォーマンスを持つ可能性が高いことは明らかであろう．したがって，定理2は，定理1にとともに提案法のメリットを示すきわめて重要な結果である．

4．北海道の稲作生産性に対する気温の影響分析

4.1 モデル

ここで，北海道における稲作の生産性を解析するのに次のモデルを用いる．

$$q_i = as_i^{h_i} \quad (i = 1, 2, \ldots, n) \tag{53}$$

ただし，q_i は i 年の収穫量（北海道全域），s_i は同年の作付面積（北海道全域）を表す．また，a は未知の定数係数であり，h_i は作付面積の単位増加率に対する収穫量の増加率，つまり作付面積に関する収穫量の弾力性係数を表す．h_i における添字 i は 弾力性係数が年によって変わる要因の影響を受けることを意味する．(53) 式の両辺の対数をとり確率変動の部分を加えれば，次式が得られる．

$$\ln\{q_i\} = \ln\{a\} + \ln\{s_i\}h_i + \varepsilon_i \tag{54}$$

ただし，ε_i は $\ln\{a\} + \ln\{s_i\}h_i$ を用いて $\ln\{q_i\}$ を表現するときの確率誤差を表す．(54) 式から解るように，h_i は対数変換後の稲作作付面積の生産性を表すものである．

言うまでもなく，肥料の使い方や品種改良などの栽培の技術的要因も直接稲作の生産性に影響を及ぼすが，自然条件として気象変数の影響が最も重要な要因と考えられる．それゆえ本稿では，稲作の生産性解析のマクロモデルとして気象変数の影響だけをモデルに取り込む．また，気象変数には気温や降水量など種々あるが，北海道の場合，気温の影響が一番大きいと考えられる．そこで，弾力性係数 h_i を次のように各月の平均気温の線形関数とする．

$$h_i = b_0 + \sum_{j=1}^{12} t_j^{(i)} b_j \tag{55}$$

ただし，$t_j^{(i)}$ は i 年 j 月の札幌における平均気温を表す．$b_0, b_1, \ldots, b_{12}$ は未知係数であり，特に b_1 から b_{12} までの係数は各月平均気温の稲作生産性への影響を表すもので，ここで，これらのパラメータを各月の気温効果と呼ぶ．(55) 式を用いると，(54) 式のモデルは次のように書き換えら

れる．

$$\ln\{q_i\} = \ln\{a\} + \ln\{s_i\}b_0 + \sum_{j=1}^{12} \ln\{s_i\}t_j^{(i)} b_j + \varepsilon_i \tag{56}$$

さらに，$y_i = \ln\{q_i\}$，$\beta_0 = \ln\{a\}$，$x_{i1} = \ln\{s_i\}$，$x_{i(j+1)} = \ln\{s_i\}t_j^{(i)}$ $(j = 1, 2, \ldots, 12)$，$\beta_{j+1} = b_j$ $(j = 0, 1, \ldots, 12)$ と置くと，(56) 式のモデルを (5) 式の線形回帰モデルで表現することができる．ここで，説明変数の数 m は 13 である．

4.2 モデルの選択

モデルの推定と選択に用いるデータは明治 23 年から平成 14 年にかける 113 年間のものである[1]．したがって，標本のサイズは $n + 1 = 113$ となる．

基本モデルにおける説明変数の数が 13 であるので，対応する係数成分の数も 13 となる．提案した係数成分制約法の手順を踏まえれば，(32) 式のモデルにおける枢軸成分の値 $z_1, z_2, \ldots, z_{13}$ と $d_1, d_2, \ldots, d_{13}$ の値を得る．定理 1 で示されているように，各枢軸成分 z_i の絶対値によって各係数成分に対する制約の優先順位が決まる．ただし，(43) 式の誤差分散最尤推定の計算に基本モデルの誤差平方和 $\sum_{i=1}^{n-m} e_i^2$ が必要となる．ここで，$\sum_{i=1}^{n-m} e_i^2 = 14.3449$ である．このとき，各係数成分のモデルにおける z_i と d_i の値および係数成分制約の優先順位は表 1 のようになる．

表 1．各係数成分のモデルにおける z_i と d_i の値及び制約の優先順位

係数成分の番号 (i)	0	1	2	3	4	5	6
z_i の値	-131.1326	16.0056	0.5562	0.4767	0.0786	0.1768	-0.5219
d_i の値	-10.63	695.00	-289.50	-196.77	-172.62	-153.49	-134.84
制約の優先順位		13	10	7	4	6	9
係数成分の番号 (i)	7	8	9	10	11	12	13
z_i の値	-0.0183	-0.7931	0.0267	-0.4831	-0.1142	0.0412	-1.7248
d_i の値	-131.83	-116.15	-108.45	-99.62	-98.40	-92.25	-4.07
制約の優先順位	1	11	2	8	5	3	12

さて，各係数成分に対する制約の優先順位にしたがい，順次各係数成分にゼロ制約を課すことによって，考え得る候補モデルを構築しよう．(43) 式，(42) 式および (41) 式を適用すれば，k 個の係数成分制約を適用したモデルの誤差分散の推定値 $\hat{\sigma}_k^2$ と AIC_k の値を計算することができる．その結果は表 2 に示されている．表 2 から解るように，最小 AIC 値を得ているモデルは，適用される係数成分制約の数が 8 のモデルであり，制約が適用される係数成分を優先順位で並べると 7, 9, 12, 4, 11, 5, 3, 10 である．他方，何の制約も付けない基本モデルについて，誤差分散の推定値は $\hat{\sigma}_0^2 = 0.12695$ であり，AIC の値は $AIC_0 = 117.45$ である．また，適用される係数成分制約数

[1] 資料は農林水産省「寒冷地農業調査」,「北海道農業累年統計表」及び「作物統計」より引用.

が 6 から 10 までの各モデルの AIC 値に明らかな変化がない．したがって，実際の応用では各方面から計算の結果を吟味してモデル選択を行った方がよいと思われる．

表2．各候補モデルの誤差分散の推定値 と AIC の値

適用係数制約数 (k)	1	2	3	4	5	6	7
追加係数制約の番号	7	9	12	4	11	5	3
誤差分散推定値 $\hat{\sigma}_k^2$	0.12695	0.12696	0.12697	0.12702	0.12714	0.12742	0.12943
自由パラメータ数	14	13	12	11	10	9	8
AIC_k の値	115.45	113.46	111.47	109.52	107.62	105.87	105.64
適用係数制約数 (k)	8	9	10	11	12	13	14
追加係数制約の番号	10	6	2	8	13	1	0
誤差分散推定値 $\hat{\sigma}_k^2$	0.13149	0.13390	0.13664	0.14221	0.16853	2.43561	154.611
自由パラメータ数	14	13	12	11	10	9	8
AIC_k の値	105.43	105.48	105.77	108.28	125.47	425.27	892.30

4.3 推定の結果

ここで，忠実に最小 AIC 法に基づいて 8 個の係数成分制約を適用したモデルを最良のモデルとして選び，このモデルについてパラメターの推定を行う．(53) 式と (55) 式で定義されるモデルのパラメータは，得られた回帰モデルのパラメータと一対一に対応するので，得られた回帰モデルに関する推定結果からそれらパラメータの推定値を求めることができる．このように得られたモデルの推定結果は次の通りである．

$$q_i = 1.7565 s_i^{h_i} \tag{57}$$

$$\begin{aligned} h_i = \ & 0.4243 + 0.0007t_1^{(i)} + 0.0021t_2^{(i)} + 0.0020t_3^{(i)} + 0.0031t_4^{(i)} + 0.0067t_5^{(i)} + 0.0088t_6^{(i)} \\ & +0.0092t_7^{(i)} + 0.0084t_8^{(i)} + 0.0024t_9^{(i)} + 0.0009t_{10}^{(i)} - 0.0023t_{11}^{(i)} + 0.0044t_{12}^{(i)} \end{aligned} \tag{58}$$

この結果は稲作生産性の分析や稲の作況予測などに重要な意味を持つものである．(58) 式のモデルから解るように，気温効果は，7 月をピークに 5 月から 8 月までの各月における数値が突出して高い．この結果は稲作の常識にも合致している．また，明らかに稲の生産と関係しない季節における気温効果は数値が相対的に小さいながらも，ほとんどプラスとなることが結果から読み取れる．これは各月気温の間における自己相関によるものと考えられる．

前述の分析のモデルにおいて，説明変数の数が 13 個もある．これについて従来の係数制約法でモデルの選択を行うことは相当困難なことが想定できる．また，従来の方法では，各月の気温効果がすべてモデルに現れることができないため，それぞれの強弱を比較できないだけでなく，特別な気温効果の除外に伴う推定値と予測値に顕著な偏りの出る可能性が高い．このような点も新しく提案した係数制約法のメリットと言えるであろう．

5．おわりに

本論文では，従来の回帰分析における説明変数の取捨によるモデルを構築する係数制約法に代わって，係数成分制約法を新しい方法として提案した．提案法では，全ての利用可能な説明変数を取り込んだ回帰モデルを基本モデルとし，その回帰係数で係数成分を構成する．そして，構成した係数成分について制約を課すことによって，候補モデルを構成し，そのうえでモデルの選択と推定を行なう．本稿の提案法を利用すれば，候補モデルの数が減少し，モデルの選択がより容易にできる他，従来の方法よりよいモデルを得ることも可能である．また，提案法を北海道の稲作生産性に対する気温の影響分析に適用することで，稲作の生産性分析や稲の作況予測に対して興味深い示唆が得られた．

参考文献

[1] 北川源四郎：「FORTRAN77 時系列解析プログラミング」，岩波書店 (1993)

[2] 北川源四郎：「時系列解析入門」，岩波書店 (2005)

[3] 小西貞則，北川源四郎：「情報量規準」，朝倉書店 (2004)

[4] 坂元慶行，石黒真木夫，北川源四郎：「情報量統計学」，共立出版 (1983)

[5] 鈴木友彦，後藤正幸，俵　信彦："線形回帰モデルのベイズ最適な予測法に関する研究"，日本経営工学会論文誌，Vol.51, No.1, pp.60–69 (2000)

[6] 松嶋敏泰："統計モデル選択の概要"，オペレーションズ・リサーチ，Vol.41, No.7, pp.369–374 (1996)

[7] Draper, N.R. and Smith, H.:"Applied Regression Analysis (2nd edition)," Wiley, New York (1981)

[8] Golub, G.H. and Loan, C.F.V.:"Matrix Computations (2nd edition)," pp.427–437, The Johns Hopkins University Press, Baltimore and London (1989)

[9] Sen, A. and Srivastava, M.:"Rgression Analysis: Theory, Methods, and Applications," Springer-Verlag, New York (1990)

付録 A：最小分散予測の原則に関する説明

係数成分 α_i に対して適当に与えられた定数 c をもってそれに

$$\alpha_i = c \tag{59}$$

という制約を付加する場合，(32) 式のモデルより，

$$z_i = cd_i + \psi_i \tag{60}$$

という関係が得られる．ψ_i の分布に関する仮定により，z_i^2 の期待値は

$$\mathrm{E}\{z_i^2\} = c^2 d_i^2 + \sigma^2 \tag{61}$$

となる．さらに，z_i の分散 $\mathrm{Var}\{z_i\}$ とその期待値 $\mathrm{E}\{z_i\}$ との関係を利用すれば，次式が得られる．

$$\begin{aligned} \mathrm{Var}\{z_i\} &= \mathrm{E}\{z_i^2\} - (\mathrm{E}\{z_i\})^2 \\ &= c^2 d_i^2 + \sigma^2 - (\mathrm{E}\{z_i\})^2 \end{aligned} \tag{62}$$

明らかなように，(62) 式において，$(\mathrm{E}\{z_i\})^2$，$d_i^2 > 0$ 及び σ^2 の何れも c に依存しないことから，z_i の分散 $\mathrm{Var}\{z_i\}$ は c^2 に比例することが解る．よって，$\mathrm{Var}\{z_i\}$ は $c = 0$ のときに最小値を取ることが自明である．したがって，各係数成分 α_i についての制約において，(34) 式のようなゼロ制約は z_i に安定性の最もよい予測値を与えるものである．

付録B：定理1の証明

ここで，(32) 式のモデルにおける枢軸ベクトル要素の系列 $z_1, z_2, \ldots, z_m$ を適当な順序で並べ替えたものを

$$\widetilde{z}_1, \widetilde{z}_2, \ldots, \widetilde{z}_m \tag{63}$$

とし，これに対応して $d_1, d_2, \ldots, d_m$ の系列と $\alpha_1, \alpha_2, \ldots, \alpha_m$ の系列を並べ替えたものをそれぞれ次のように表す．

$$\widetilde{d}_1, \widetilde{d}_2, \quad \ldots, \quad \widetilde{d}_m \tag{64}$$

$$\widetilde{\alpha}_1, \widetilde{\alpha}_2, \quad \ldots, \quad \widetilde{\alpha}_m \tag{65}$$

また，$0 < k < m$ を満たす整数 k について (37) 式で定義される係数成分制約集合 $\boldsymbol{\mathcal{A}}$ の部分集合 $\widetilde{\boldsymbol{\mathcal{A}}}_k$ を $k = 0$ のときに空集合 $\widetilde{\boldsymbol{\mathcal{A}}}_0 = \Phi$ で定義し，$k > 0$ のときに

$$\widetilde{\boldsymbol{\mathcal{A}}}_k \equiv \{\mathrm{Re}(\widetilde{\alpha}_{m-k+1}), \mathrm{Re}(\widetilde{\alpha}_{m-1}), \ldots, \mathrm{Re}(\widetilde{\alpha}_m)\} \tag{66}$$

で定義する．ただし，$\mathrm{Re}(\widetilde{\alpha}_i)$ は (36) 式で定義されている．

第 3.2 節と第 3.3 節におけるモデルの中心化と独立成分化から明らかなように，尤度関数の定義において，(9) 式の基本モデルは，(15) 式，(30) 式と (17) 式のモデルと同値である．そこで，(20) 式と (22) 式及び $\boldsymbol{H}_0\boldsymbol{\varepsilon}$ と $\boldsymbol{H}_2\boldsymbol{\varepsilon}$ との独立性により，(15) 式と (17) 式の部分モデルに対応する尤度関数は

$$L_1(\alpha_0, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\{-\frac{1}{2\sigma^2}(z_0 - d_0\alpha_0)^2\} \times (2\pi\sigma^2)^{-\frac{n-m}{2}} \exp\{-\frac{1}{2\sigma^2}\boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}\} \tag{67}$$

で表現される．また，(31) 式により (30) 式の部分モデルに対応する尤度関数は

$$L_2(\boldsymbol{\alpha},\sigma^2)=(2\pi\sigma^2)^{-\frac{m}{2}}\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{m}(\widetilde{z}_i-\widetilde{d}_i\widetilde{\alpha}_i)^2\} \tag{68}$$

となる．さらに，各部分モデル間の独立性により，基本モデルの尤度関数は

$$\begin{aligned} L(\alpha_0,\boldsymbol{\alpha},\sigma^2) &= L_1(\alpha_0,\sigma^2)\times L_2(\boldsymbol{\alpha},\sigma^2) \\ &= (2\pi\sigma^2)^{-\frac{n+1}{2}}\exp\Big[-\frac{1}{2\sigma^2}\{\sum_{i=0}^{m}(\widetilde{z}_i-\widetilde{d}_i\widetilde{\alpha}_i)^2+\boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}\}\Big] \end{aligned} \tag{69}$$

となり，対数尤度は

$$\begin{aligned} \ell(\alpha_0,\boldsymbol{\alpha},\sigma^2) &= \ln\{L(\alpha_0,\boldsymbol{\alpha},\sigma^2)\} \\ &= -\frac{n+1}{2}\ln\{2\pi\sigma^2\}-\frac{1}{2\sigma^2}\{\sum_{i=0}^{m}(\widetilde{z}_i-\widetilde{d}_i\widetilde{\alpha}_i)^2+\boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}\} \end{aligned} \tag{70}$$

となる．ただし，パラメータの最尤推定は，文字通り尤度関数または対数尤度の最大化によって得られるため，平均期待 α_0 を未知パラメータとして取り扱うとき，一定の σ^2 の値に対してその最尤推定値 $\widehat{\alpha}_0$ は $\alpha_0=\widehat{\alpha}_0$ および

$$z_0-d_0\alpha_0=0 \tag{71}$$

によって (33) 式で得られる．また，基本モデルに $\widetilde{\boldsymbol{\mathcal{A}}}_k$ における係数成分制約を適用したモデルでは，対応する係数成分の最尤推定は

$$\widetilde{z}_i-\widetilde{d}_i\widetilde{\alpha}_i=0 \quad (i=1,2,\ldots,m-k) \tag{72}$$

によって得られ，σ^2 の最尤推定値は

$$\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}}_k)=\frac{1}{n+1}(\sum_{i=m-k+1}^{m}\widetilde{z}_i^2+\boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}) \tag{73}$$

で与えられる．また，$\widetilde{\boldsymbol{\mathcal{A}}}_k$ における係数成分制約が適用された係数成分ベクトル $\boldsymbol{\alpha}(\widetilde{\boldsymbol{\mathcal{A}}}_k)$ の最尤推定を $\widehat{\boldsymbol{\alpha}}(\widetilde{\boldsymbol{\mathcal{A}}}_k)$ とすると，最大対数尤度は

$$\ell(\widehat{\alpha}_0,\widehat{\boldsymbol{\alpha}}(\widetilde{\boldsymbol{\mathcal{A}}}_k),\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}}_k))=-\frac{n+1}{2}\Big[\ln\{2\pi\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}}_k)\}+1\Big] \tag{74}$$

で与えられる．基本モデルにおけるパラメータ数は $m+2$ であるが，$\widetilde{\boldsymbol{\mathcal{A}}}_k$ における k 個の係数成分制約を適用しているので，自由パラメータ数は $m-k+2$ となる．したがって，AIC の定義により，このモデルの AIC 値は次のように与えられる（坂元・石黒・北川 (1983)）．

$$\begin{aligned} AIC(\widetilde{\boldsymbol{\mathcal{A}}}_k)) &= -2\ell(\widehat{\alpha}_0,\widehat{\boldsymbol{\alpha}}(\widetilde{\boldsymbol{\mathcal{A}}}_k),\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}}_k))+2(m-k+2) \\ &= (n+1)\ln\{\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}}_k)\}+\lambda \end{aligned} \tag{75}$$

ただし，λ は (42) 式で定義される．

明らかなように，(75) 式で定義した AIC の値は $\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}_k})$ の単調増加関数であるため，適用される係数成分制約数 k が一定のとき，より小さい $\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}_k})$ の値はより小さい AIC 値を導く．最小 AIC 法によれば，可能な限り $\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}_k})$ の値を小さくし，AIC の最小値を与える係数成分制約を適用した方が望ましい．(73) 式から解るように，集合 $\{z_1^2, z_2^2, \ldots, z_m^2\}$ の最も小さい k 個の要素で $\tilde{z}_1^2, \tilde{z}_2^2, \ldots, \tilde{z}_k^2$ に充てば，$\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}_k})$ が最小となる．したがって，(63) 式を (38) 式で取り替え，さらに，$\widetilde{\boldsymbol{\mathcal{A}}_k}$ を (40) 式で定義される $\boldsymbol{\mathcal{A}}_k$ で取り替えれば最小 AIC を導くモデルが得られる．このようにして，(73) 式と (75) 式で定義される $\widehat{\sigma}^2(\widetilde{\boldsymbol{\mathcal{A}}_k})$ と $AIC(\widetilde{\boldsymbol{\mathcal{A}}_k})$ はそれぞれ (43) 式と (41) 式の定義に一致する．よって，定理1が証明された．

付録C：補題2の証明

行列 $\boldsymbol{V}$ が直交行列のことを利用すれば，(46) 式は次のように書き換えられる．

$$\begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{V}_{11}^{\mathrm{t}} & \boldsymbol{V}_{21}^{\mathrm{t}} \\ \boldsymbol{V}_{12}^{\mathrm{t}} & \boldsymbol{V}_{22}^{\mathrm{t}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{bmatrix} \tag{76}$$

(76) 式に次の関係が含まれる．

$$\boldsymbol{\beta}_2 = \boldsymbol{V}_{12}^{\mathrm{t}}\boldsymbol{\alpha}_1 + \boldsymbol{V}_{22}^{\mathrm{t}}\boldsymbol{\alpha}_2 \tag{77}$$

この関係に (4 7) 式の係数制約を用いると，

$$\boldsymbol{\alpha}_2 = -\boldsymbol{V}_{22}^{-\mathrm{t}}\boldsymbol{V}_{12}^{\mathrm{t}}\boldsymbol{\alpha}_1 \tag{78}$$

と得られる．これは常に保たなければならないので，$\boldsymbol{\alpha}_1$ と $\boldsymbol{\alpha}_2$ の最尤推定値に対しても

$$\widehat{\boldsymbol{\alpha}}_2 = -\boldsymbol{V}_{22}^{-\mathrm{t}}\boldsymbol{V}_{12}^{\mathrm{t}}\widehat{\boldsymbol{\alpha}}_1 \tag{79}$$

のように成立する．したがって，$\widehat{\boldsymbol{\alpha}}_2$ の分散共分散行列を $\mathrm{Var}\{\widehat{\boldsymbol{\alpha}}_2\} = \mathrm{E}\{(\widehat{\boldsymbol{\alpha}}_2 - \boldsymbol{\alpha}_2)(\widehat{\boldsymbol{\alpha}}_2 - \boldsymbol{\alpha}_2)^{\mathrm{t}}\}$ で，$\widehat{\boldsymbol{\alpha}}_1$ と $\widehat{\boldsymbol{\alpha}}_2$ との共分散行列を $\mathrm{Cov}\{\widehat{\boldsymbol{\alpha}}_1, \widehat{\boldsymbol{\alpha}}_2\} = \mathrm{E}\{(\widehat{\boldsymbol{\alpha}}_1 - \boldsymbol{\alpha}_1)(\widehat{\boldsymbol{\alpha}}_2 - \boldsymbol{\alpha}_2)^{\mathrm{t}}\}$ で定義すれば，

$$\mathrm{Var}\{\widehat{\boldsymbol{\alpha}}_2\} = -\boldsymbol{V}_{22}^{-\mathrm{t}}\boldsymbol{V}_{12}^{\mathrm{t}}\mathrm{Cov}\{\widehat{\boldsymbol{\alpha}}_1, \widehat{\boldsymbol{\alpha}}_2\} \tag{80}$$

という関係がある．補題1によると，基本モデルからの結果と矛盾しないために，$\mathrm{Cov}\{\widehat{\boldsymbol{\alpha}}_1, \widehat{\boldsymbol{\alpha}}_2\} = \boldsymbol{O}$ と設定しなければならない．こうして，(80) 式から

$$Var\{\widehat{\boldsymbol{\alpha}}_2\} = \boldsymbol{O} \tag{81}$$

と得られる．これは，任意の標本データに対して，$\widehat{\boldsymbol{\alpha}}_2 = \boldsymbol{\alpha}_2$ は常に定数であることを意味するから，それに定数の制約を付加する必要がある．前述のように，(49) 式を置けば最小分散予測が得られる．よって，補題2が証明された．

付録D：定理２の証明

付録B及び第 3.5 節における問題の設定から明らかなように，(15) 式，(49) 式と (17) 式のモデルの尤度関数は

$$\begin{aligned} L(\alpha_0, \boldsymbol{\beta}_1, \sigma^2) &= (2\pi\sigma^2)^{-\frac{n+1}{2}} \exp\Big[-\frac{1}{2\sigma^2}\{(z_0 - d_0\alpha_0)^2 \\ &\quad +(\boldsymbol{w}_1 - \boldsymbol{T}_{11}\boldsymbol{\beta}_1)^{\mathrm{t}}(\boldsymbol{w}_1 - \boldsymbol{T}_{11}\boldsymbol{\beta}_1) + \boldsymbol{w}_2^{\mathrm{t}}\boldsymbol{w}_2 + \boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}\}\Big] \end{aligned} \tag{82}$$

で表され，対数尤度は

$$\begin{aligned} \ell(\alpha_0, \boldsymbol{\alpha}, \sigma^2) &= -\frac{n+1}{2}\ln\{2\pi\sigma^2\} - \frac{1}{2\sigma^2}\{(z_0 - d_0\alpha_0)^2 \\ &\quad +(\boldsymbol{w}_1 - \boldsymbol{T}_{11}\boldsymbol{\beta}_1)^{\mathrm{t}}(\boldsymbol{w}_1 - \boldsymbol{T}_{11}\boldsymbol{\beta}_1) + \boldsymbol{w}_2^{\mathrm{t}}\boldsymbol{w}_2 + \boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}\} \end{aligned} \tag{83}$$

となる．一定の σ^2 の値に対して，パラメータ α_0 と $\boldsymbol{\beta}_1$ の最尤推定値はそれぞれ (71) 式と

$$\boldsymbol{w}_1 - \boldsymbol{T}_{11}\boldsymbol{\beta}_1 = \boldsymbol{0}_{m-k} \tag{84}$$

で得られる．そして，σ^2 の最尤推定値は

$$\widehat{\sigma}_1^2(k) = \frac{1}{n+1}(\boldsymbol{w}_2^{\mathrm{t}}\boldsymbol{w}_2 + \boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}) \tag{85}$$

で与えられる．一方，(15) 式，(50) 式と (17) 式のモデルの尤度関数は

$$\begin{aligned} L(\alpha_0, \boldsymbol{\alpha}_1, \sigma^2) &= (2\pi\sigma^2)^{-\frac{n+1}{2}} \exp\Big[-\frac{1}{2\sigma^2}\{(z_0 - d_0\alpha_0)^2 \\ &\quad +(\boldsymbol{z}_1 - \boldsymbol{D}_1\boldsymbol{\alpha}_1)^{\mathrm{t}}(\boldsymbol{z}_1 - \boldsymbol{D}_1\boldsymbol{\alpha}_1) + \boldsymbol{z}_2^{\mathrm{t}}\boldsymbol{z}_2 + \boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}\}\Big] \end{aligned} \tag{86}$$

で与えられ，対数尤度は

$$\begin{aligned} \ell(\alpha_0, \boldsymbol{\alpha}, \sigma^2) &= -\frac{n+1}{2}\ln\{2\pi\sigma^2\} - \frac{1}{2\sigma^2}\{(z_0 - d_0\alpha_0)^2 \\ &\quad +(\boldsymbol{z}_1 - \boldsymbol{D}_1\boldsymbol{\alpha}_1)^{\mathrm{t}}(\boldsymbol{z}_1 - \boldsymbol{D}_1\boldsymbol{\alpha}_1) + \boldsymbol{z}_2^{\mathrm{t}}\boldsymbol{z}_2 + \boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}\} \end{aligned} \tag{87}$$

となる．同じように，パラメータ α_0 と $\boldsymbol{\alpha}_1$ の最尤推定値はそれぞれ (71) 式と

$$\boldsymbol{z}_1 - \boldsymbol{D}_1\boldsymbol{\alpha}_1 = \boldsymbol{0}_{m-k} \tag{88}$$

によって求められ，σ^2 の最尤推定値は

$$\widehat{\sigma}_2^2(k) = \frac{1}{n+1}(\boldsymbol{z}_2^{\mathrm{t}}\boldsymbol{z}_2 + \boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}) \tag{89}$$

で得られる．さらに，(17) 式と (22) 式により，$\boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}$ の基本モデルに関する期待値は

$$\mathrm{E}\{\boldsymbol{e}^{\mathrm{t}}\boldsymbol{e}\} = (n-m)\sigma^2 \tag{90}$$

のように得られる．また，(44) 式と (21) 式により，$\boldsymbol{w}_2^{\mathrm{t}}\boldsymbol{w}_2$ の基本モデルに関する期待値は

$$\mathrm{E}\{\boldsymbol{w}_2^{\mathrm{t}}\boldsymbol{w}_2\} = \boldsymbol{\beta}_2^{\mathrm{t}}\boldsymbol{T}_{22}^{\mathrm{t}}\boldsymbol{T}_{22}\boldsymbol{\beta}_2 + k\sigma^2 \tag{91}$$

で与えられ，(45) 式と (31) 式により，$\boldsymbol{z}_2^{\mathrm{t}}\boldsymbol{z}_2$ の基本モデルに関する期待値は

$$\mathrm{E}\{\boldsymbol{z}_2^{\mathrm{t}}\boldsymbol{z}_2\} = \boldsymbol{\alpha}_2^{\mathrm{t}}\boldsymbol{D}_2^2\boldsymbol{\alpha}_2 + k\sigma^2 \tag{92}$$

で得られる．よって，基本モデルに関する $\hat{\sigma}_1^2(k)$ と $\hat{\sigma}_2^2(k)$ の期待値はそれぞれ次のように求められる．

$$\mathrm{E}\{\hat{\sigma}_1^2(k)\} = \frac{1}{n+1}\{\boldsymbol{\beta}_2^{\mathrm{t}}\boldsymbol{T}_{22}^{\mathrm{t}}\boldsymbol{T}_{22}\boldsymbol{\beta}_2 + (n-m+k)\sigma^2\} \tag{93}$$

$$\mathrm{E}\{\hat{\sigma}_2^2(k)\} = \frac{1}{n+1}\{\boldsymbol{\alpha}_2^{\mathrm{t}}\boldsymbol{D}_2^2\boldsymbol{\alpha}_2 + (n-m+k)\sigma^2\} \tag{94}$$

補題２により，(47) 式が成り立つ場合，(49) 式は必ず成り立つので

$$\boldsymbol{\beta}_2^{\mathrm{t}}\boldsymbol{T}_{22}^{\mathrm{t}}\boldsymbol{T}_{22}\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2^{\mathrm{t}}\boldsymbol{D}_2\boldsymbol{\alpha}_2 = 0 \tag{95}$$

という関係が得られる．しかし，逆に (49) 式が成り立つ場合，(47) 式は必ずしも満たされない．したがって，次の関係が成立する．

$$\boldsymbol{\beta}_2^{\mathrm{t}}\boldsymbol{T}_{22}^{\mathrm{t}}\boldsymbol{T}_{22}\boldsymbol{\beta}_2 \geq \boldsymbol{\alpha}_2^{\mathrm{t}}\boldsymbol{D}_2\boldsymbol{\alpha}_2 = 0 \tag{96}$$

つまり，(47) 式と (49) 式の少なくとも一つが成り立てば，(51) 式の不等式が必ず成り立つ．してがって，定理２が証明された．